

D. Srivastava

1653

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/808,031A

DATE: 05/06/1999
TIME: 08:25:11

INPUT SET: S31745.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

#30

SEQUENCE LISTING

ENTERED

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

(1) General Information:

(i) APPLICANT: Inouye, Sumiko
Hsu, Mei-Yin
Eagle, Susan
Inouye, Masayori

(ii) TITLE OF INVENTION: PROKARYOTIC REVERSE TRANSCRIPTASE

(iii) NUMBER OF SEQUENCES: 52

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: WEISER & ASSOCIATES
(B) STREET: 230 South Fifteenth Street, Suite 500
(C) CITY: Philadelphia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19102

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/808,031
(B) FILING DATE: 03-MAR-1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Weiser, Gerard J.
(B) REGISTRATION NUMBER: 19,763
(C) REFERENCE/DOCKET NUMBER: 377(913).5888P

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 215-875-8383
(B) TELEFAX: 215-875-8394

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2176 base pairs

INPUT SET: S31745.raw

```

47          (B) TYPE: nucleic acid
48          (C) STRANDEDNESS: double
49          (D) TOPOLOGY: linear
50
51      (ii) MOLECULE TYPE: cDNA
52
53
54      (ix) FEATURE:
55          (A) NAME/KEY: CDS
56          (B) LOCATION: 640..2094
57
58
59      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
60
61      TCATCCGCGC GGACACCCCC TCCTACGTGC CCCCCGACGC GGAGAGCGGC GTGGAGACGG      60
62
63      TGTACCGCGT TTCCCTGGAT GGTCACCTGG TGGCGGTGGA GTGGGGCCCG CGCACGGGCT      120
64
65      CGCCGCGTCA CCAGCGGCTC TGGTTCGACT CGGATGCGGA AGCCCCCGGA GCCTACTTCG      180
66
67      CGCGCCTCGA GAAGTTGGCG GCTGACGGCT ACATCGACGC GGCCTCGGCA TTGGTCTAAA      240
68
69      CCCTTCAACC ACGGCTCGGC CGCCACGCGC GGCCGGCAGG ACAGGTGCGA CGAACAGACG      300
70
71      ACGACGTGCG CTTACGCGC GAGCAGCCGA GAGAGGTCCG GAGTGCATCA GCCTGAGCGC      360
72
73      CTCGAGCGGC GGAGCGGCGT TGCGCCGCTC CGGTTGGAAT GCAGGACACT CTCCGCAAGG      420
74
75      TAGCCTGTTC TTGGCTCTCT CCCTCCTAGG CACTACGGCC AGGGTGGGTA GCGGAGCCAA      480
76
77      CGACGCCACC GCCGTTTACC CACCCCGGCC GTAGTGCCTA GGAGGGGAGA GCCGGTGAGG      540
78
79      CTACCGTGCC CCAGGTAAGA TGGTGGTGCT TTCCCGGCCCT CCGTCGACTG CTCGCGCCAT      600
80
81      GTCCCGTCTT CCATCGCCGC GCCCGCCCAA GGTGCAGAC ATG ACC GCC AGG CTG      654
82                                     Met Thr Ala Arg Leu
83                                     1 5
84
85      GAC CCG TTC GTC CCC GCA GCT TCG CCG CAG GCC GTG CCC ACG CCC GAG      702
86      Asp Pro Phe Val Pro Ala Ala Ser Pro Gln Ala Val Pro Thr Pro Glu
87                                     10 15 20
88
89      CTC ACC GCT CCG TCG TCA GAC GCG GCC GCG AAG CGT GAA GCC CGC CGG      750
90      Leu Thr Ala Pro Ser Ser Asp Ala Ala Ala Lys Arg Glu Ala Arg Arg
91                                     25 30 35
92
93      CTC GCG CAC GAA GCG TTG CTC GTC CGC GCG AAG GCC ATC GAC GAA GCG      798
94      Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys Ala Ile Asp Glu Ala
95                                     40 45 50
96
97      GGC GGC GCC GAC GAC TGG GTG CAG GCG CAG CTC GTC TCC AAG GGG CTC      846
98      Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu Val Ser Lys Gly Leu
99      55 60 65

```

INPUT SET: S31745.raw

100																		894
101	GCG	GTC	GAG	GAC	CTG	GAC	TTC	TCC	AGC	GCC	TCC	GAG	AAG	GAC	AAG	AAG		
102	Ala	Val	Glu	Asp	Leu	Asp	Phe	Ser	Ser	Ala	Ser	Glu	Lys	Asp	Lys	Lys		
103	.70					.75					.80					.85		
104																		
105	GCC	TGG	AAG	GAG	AAG	AAG	AAG	GCC	GAG	GCC	ACC	GAG	CGC	CGC	GCG	CTG		942
106	Ala	Trp	Lys	Glu	Lys	Lys	Lys	Ala	Glu	Ala	Thr	Glu	Arg	Arg	Ala	Leu		
107					.90					.95					.100			
108																		
109	AAG	CGT	CAG	GCG	CAC	GAG	GCG	TGG	AAG	GCC	ACG	CAC	GTG	GGC	CAC	CTG		990
110	Lys	Arg	Gln	Ala	His	Glu	Ala	Trp	Lys	Ala	Thr	His	Val	Gly	His	Leu		
111				.105					.110					.115				
112																		
113	GGC	GCG	GGC	GTG	CAC	TGG	GCG	GAG	GAC	CGC	CTG	GCC	GAC	GCG	TTC	GAC		1038
114	Gly	Ala	Gly	Val	His	Trp	Ala	Glu	Asp	Arg	Leu	Ala	Asp	Ala	Phe	Asp		
115			.120					.125					.130					
116																		
117	GTG	CCC	CAC	CGC	GAG	GAG	CGC	GCC	CGG	GCC	AAC	GGC	CTG	ACG	GAG	CTG		1086
118	Val	Pro	His	Arg	Glu	Glu	Arg	Ala	Arg	Ala	Asn	Gly	Leu	Thr	Glu	Leu		
119		.135					.140					.145						
120																		
121	GAC	TCC	GCG	GAG	GCG	CTG	GCC	AAG	GCG	CTG	GGG	CTG	AGC	GTC	TCC	AAG		1134
122	Asp	Ser	Ala	Glu	Ala	Leu	Ala	Lys	Ala	Leu	Gly	Leu	Ser	Val	Ser	Lys		
123	.150					.155					.160					.165		
124																		
125	CTC	CGC	TGG	TTC	GCG	TTC	CAC	CGG	GAG	GTC	GAC	ACG	GCC	ACG	CAC	TAC		1182
126	Leu	Arg	Trp	Phe	Ala	Phe	His	Arg	Glu	Val	Asp	Thr	Ala	Thr	His	Tyr		
127					.170					.175					.180			
128																		
129	GTG	AGC	TGG	ACC	ATT	CCG	AAG	CGG	GAC	GGC	AGC	AAG	CGC	ACG	ATT	ACG		1230
130	Val	Ser	Trp	Thr	Ile	Pro	Lys	Arg	Asp	Gly	Ser	Lys	Arg	Thr	Ile	Thr		
131				.185					.190					.195				
132																		
133	TCC	CCC	AAG	CCT	GAG	CTG	AAG	GCA	GCG	CAG	CGC	TGG	GTG	CTG	TCC	AAC		1278
134	Ser	Pro	Lys	Pro	Glu	Leu	Lys	Ala	Ala	Gln	Arg	Trp	Val	Leu	Ser	Asn		
135			.200					.205					.210					
136																		
137	GTC	GTG	GAG	CGG	CTG	CCG	GTC	CAC	GGC	GCC	GCC	CAC	GGC	TTC	GTG	GCG		

RAW SEQUENCE LISTING PATENT APPLICATION US/08/808,031A

DATE: 05/06/1999
TIME: 08:25:12

INPUT SET: S31745.raw

153	ACG CTG CTG TCC CTC CTC TCC ACG GAA GCG CCG CGG GAG GCG GTC CAG	1518
154	Thr Leu Leu Ser Leu Leu Ser Thr Glu Ala Pro Arg Glu Ala Val Gln	
155	280 285 290	
156		
157	TTC CGC GGC AAG CTC CTG CAC GTC GCC AAG GGC CCG CGC GCC CTG CCC	1566
158	Phe Arg Gly Lys Leu Leu His Val Ala Lys Gly Pro Arg Ala Leu Pro	
159	295 300 305	
160		
161	CAG GGC GCC CCC ACG TCG CCC GGC ATC ACC AAC GCG CTC TGC CTG AAG	1614
162	Gln Gly Ala Pro Thr Ser Pro Gly Ile Thr Asn Ala Leu Cys Leu Lys	
163	310 315 320 325	
164		
165	CTC GAC AAG CGG CTG TCC GCC CTC GCG AAG CGG CTG GGC TTC ACC TAC	1662
166	Leu Asp Lys Arg Leu Ser Ala Leu Ala Lys Arg Leu Gly Phe Thr Tyr	
167	330 335 340	
168		
169	ACG CGC TAC GCG GAC GAC CTG ACC TTC TCC TGG ACG AAG GCG AAG CAG	1710
170	Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser Trp Thr Lys Ala Lys Gln	
171	345 350 355	
172		
173	CCC AAG CCG CGG CGG ACG CAG CGT CCC CCC GTC GCG GTC CTC CTG TCT	1758
174	Pro Lys Pro Arg Arg Thr Gln Arg Pro Pro Val Ala Val Leu Leu Ser	
175	360 365 370	
176		
177	CGC GTC CAG GAA GTG GTG GAG GCG GAG GGC TTC CGC GTG CAC CCG GAC	1806
178	Arg Val Gln Glu Val Val Glu Ala Glu Gly Phe Arg Val His Pro Asp	
179	375 380 385	
180		
181	AAG ACG CGC GTC GCC CGC AAG GGC ACG CGG CAG CGG GTC ACC GGG CTC	1854
182	Lys Thr Arg Val Ala Arg Lys Gly Thr Arg Gln Arg Val Thr Gly Leu	
183	390 395 400 405	
184		
185	GTC GTG AAT GCG GCG GGC AAG GAC GCG CCC GCG GCC CGA GTC CCG CGC	1902
186	Val Val Asn Ala Ala Gly Lys Asp Ala Pro Ala Ala Arg Val Pro Arg	
187	410 415 420	
188		
189	GAC GTC GTC CGC CAG CTC CGC GCC GCC ATC CAC AAC CGG AAG AAG GGC	1950
190	Asp Val Val Arg Gln Leu Arg Ala Ala Ile His Asn Arg Lys Lys Gly	
191	425 430 435	
192		
193	AAG CCG GGC CGC GAG GGC GAG TCG CTC GAG CAG CTC AAG GGC ATG GCC	1998
194	Lys Pro Gly Arg Glu Gly Glu Ser Leu Glu Gln Leu Lys Gly Met Ala	
195	440 445 450	
196		
197	GCC TTC ATC CAC ATG ACG GAC CCG GCC AAG GGC CGC GCC TTC CTG GCT	2046
198	Ala Phe Ile His Met Thr Asp Pro Ala Lys Gly Arg Ala Phe Leu Ala	
199	455 460 465	
200		
201	CAG CTC ACG GAG CTC GAG TCC ACG GCG AGC GCC GCT CCG CAG GCG GAG	2094
202	Gln Leu Thr Glu Leu Glu Ser Thr Ala Ser Ala Ala Pro Gln Ala Glu	
203	470 475 480 485	
204		
205	TGACGCTCAG CGCGCGTCCG TCGCCGACGT GCCGCGCGCC AGCAACGCCG CATTAGCAA	2154

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/808,031ADATE: 05/06/1999
TIME: 08:25:13

INPUT SET: S31745.raw

206
207 CTCCGTCAGC CGGCGCGGGT AC 2176
208
209
210 (2) INFORMATION FOR SEQ ID NO:2:
211
212 (i) SEQUENCE CHARACTERISTICS:
213 (A) LENGTH: 485 amino acids
214 (B) TYPE: amino acid
215 (D) TOPOLOGY: linear
216
217 (ii) MOLECULE TYPE: protein
218
219 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
220
221 Met Thr Ala Arg Leu Asp Pro Phe Val Pro Ala Ala Ser Pro Gln Ala
222 1 5 10 15
223
224 Val Pro Thr Pro Glu Leu Thr Ala Pro Ser Ser Asp Ala Ala Ala Lys
225 20 25 30
226
227 Arg Glu Ala Arg Arg Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys
228 35 40 45
229
230 Ala Ile Asp Glu Ala Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu
231 50 55 60
232
233 Val Ser Lys Gly Leu Ala Val Glu Asp Leu Asp Phe Ser Ser Ala Ser
234 65 70 75 80
235
236 Glu Lys Asp Lys Lys Ala Trp Lys Glu Lys Lys Lys Ala Glu Ala Thr
237 85 90 95
238
239 Glu Arg Arg Ala Leu Lys Arg Gln Ala His Glu Ala Trp Lys Ala Thr
240 100 105 110
241
242 His Val Gly His Leu Gly Ala Gly Val His Trp Ala Glu Asp Arg Leu
243 115 120 125
244
245 Ala Asp Ala Phe Asp Val Pro His Arg Glu Glu Arg Ala Arg Ala Asn
246 130 135 140
247
248 Gly Leu Thr Glu Leu Asp Ser Ala Glu Ala Leu Ala Lys Ala Leu Gly
249 145 150 155 160
250
251 Leu Ser Val Ser Lys Leu Arg Trp Phe Ala Phe His Arg Glu Val Asp
252 165 170 175
253
254 Thr Ala Thr His Tyr Val Ser Trp Thr Ile Pro Lys Arg Asp Gly Ser
255 180 185 190
256
257 Lys Arg Thr Ile Thr Ser Pro Lys Pro Glu Leu Lys Ala Ala Gln Arg
258 195 200 205

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/808,031A

DATE: 05/06/1999
TIME: 08:25:13

INPUT SET: S31745.raw

Line

Error

Original Text